Abnormality Detection in a Mammogram using Convolution Neural Network

Shainila Mulla[1] Shrunali Salian [2]Mathew Jose [3] Pranav Medhi[4]

Department of Computer Engineering

D.B.I.T, Mumbai, India.

Abstract— In today's time, breast cancer detection is prevalent. The time constraints imposed on radiologists is quite high and this severely impedes communication with direct care physicians and resulting in long and deleterious time-to treatment periods for patients. The purpose here is to mitigate the issues by a computer-aided system by a direct channel of communication between the patients and the doctor by applying convolution neural network, a deep learning technique which will help in classifying and layering the heterogeneous images given a sufficiently large dataset. This research pertains to designing a system to visualize the results more accurately and quickly thus streamlining the higher risk patients to get the immediate treatment , overall impacting the national standard of care.

I. INTRODUCTION

Breast cancer is the most common cancer in females, worldwide. The mortality rate is still higher than cervix cancer, even though various early diagnosis methods and suitable therapies are available for the treatment of breast cancer. As far as Indian women are concerned, most of them are not aware about the diagnosis, treatment, symptoms and causes of breast cancer. Breast cancer is referred to as one disease but there are up to 21 histological sub-categories [1]. Although no single trigger can be identified for breast cancer, certain risk factors exist, that increase a womans chance of developing it: age, family history, previous breast cancer,family history of ovar- ian cancer,age of pregnancy ,age of menstruation,entering menopause later (over age 55) increases breast cancer risks, radiation treatment to the chest, especially before 30 years of age, hormone replacement therapy,oral contraceptives increase risks slightly, if used over many years, obesity with excess caloric and fat intake and recent research suggests that women who start smoking regularly within 5 years of the onset of their menstrual periods are 70 percent more likely to develop breast cancer before the age of 50 than non-smokers [14]. However, if detected early there is 90 percent chance of being cured as it takes 5 years for a breast tumor to reach 1mm , 2 years longer to reach 5mm and one or two years to measure 2cm [8]. The datasets used by various groups in this study are MIAS, DDSM, INbreast, IRMA and BCDR.Date has been

split into training and testing majorly, however some studies have also incorporated data for validation. The percentage of splitting the data into the 2 or 3 categories respectively is randomly and varies as 80 percent, 20 percent or 70 percent, 20 percent , 10 percent.

II. PROBLEM ANALYSIS

Breast cancer is a major disease of the 21st century around the world. There is an approximately 15 percent rise in the number of new breast cancer cases registered every year in India. The 2012 mortality rate caused due to this disease was approximately 92.6 in India. However, early detection of breast cancer can save lives. The average cost for an MRI mammogram in Mumbai is Rupees 4000. Thereafter, the consulting cost of an average doctor is 1000 rupees. Thus cost is the major hindrance in the early detection of breast cancer. The most common symptom of breast cancer is a new lump or mass. A painless, hard mass that has irregular edges is more likely to be cancer, but breast cancers can be tender, soft, or rounded [17]. Additionally, there is no public website where a patient can check their mammograms online.

III. CONVOLUTION NEURAL NETWORK

Convolutional neural network primarily consists of three layers, which makes it different from the general neural network architecture.CNN is essentially used to train and find patterns in images. Each neuron in a neural network is the dot product of the weight to local region and input volume [1].Generally any neural network consists of three layers namely input,hidden ,output layers.While the three essential layers in CNN are convolution layer, pooling layer and fully-connected layer [2].

Convolution is performed by entering required values in a matrix that are obtained from the neighbouring pixels, to the pixel under examination. This matrix, is then applied to each pixel in an image. For each pixel in an image, the matrix multiplies (dot product) the pixel and its neighbouring pixels that the matrix covers by their respective matrix values. An aggregate result of all the dot products is calculated and this value is set as the pixel value in the final convolved image at the initial pixels location. As a result of convolving, the image is filtered for specific patterns or desired features which is under consideration, to enhance a particular domain of an image.The Convolving step plays an important role in image feature identification by a secured neural network.

A. Convolution Layer

Convolutional layer is the core part of the Convolutional neural network, which has local connections and weights of shared characteristics. The main function of the Convolutional layer is to learn feature representations of the inputs. As shown in above,the Convolutional layer consists of several feature maps. Each neuron of the same feature map is used to extract local characteristics of different positions in the former layer, but for single neurons, its extraction is local characteristics of same positions in former different feature map. In order to obtain a new feature, the input feature maps are first convolved with a learned kernel and then the results are passed into a nonlinear activation function. We will get different feature maps by applying different kernels. The typical activation function are sigmoid, tanh and Relu[3].
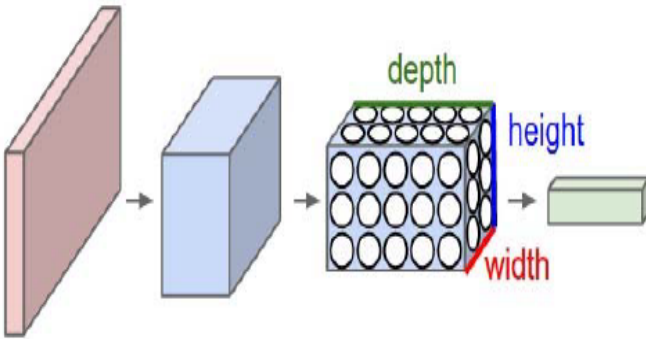

Fig. 1. A convolution neural network neuron arrangement [2]

B. Pooling Layer

The sampling system is equal to fuzzy filtering. The pooling layer has the effect of the secondary feature extraction, it may lessen the scale of the feature maps and boom the robustness of characteristic extraction. it is also positioned between two Convolutional layers. the size of feature maps in pooling layer is decided in line with the shifting step of kernels. the everyday pooling operations are common Pooling and max pooling. we can extract the high level characteristics of inputs by means of stacking several Convolutional layer and pooling layer[3].

C. Fully Connected Layer

In fashionable, the classifier of Convolutional neural community is one or extra absolutely-linked layers. They take all neurons within the preceding layer and connect them to every single neuron of modern-day layer. There is no spatial records preserved in completely-connected layers. The closing absolutely-linked layer is followed by means of an output layer. For classification duties, softmax regression is generally used because of it producing a nicely-finished chance distribution of the outputs. Any other typically used technique is SVM, which could be blended with CNNs to resolve exceptional classification responsibilities[3].

IV. IMAGE ACQUISITION

First-class tuning a CNN version calls for a large dataset for schooling and testing functions. a number of databases are available on-line. Some of them are given under:

- Mammography imaging analysis society The database is to be had on-line on the internet site http://peipa.essex.ac.united kingdom/information/mias.html for achieving the goal of education reasearch. The database incorporates 161 pairs of mediolateral indirect (MLO) for viewing photographs with 1024 pixel per inch decision. After digitizing the images, they had been annotated based entirely on overall tissue structure,which would contain fat glands,abnormal tissues like tumor, severity of abnormality (benign or malignant) by means of professional radiologists[9].
- Digital database for screening mammography It's far some other series of mammograms together with 2620 instances and forty three volumes. It is freely to be had on http://marathon.csee.usf.edu/Mammography/Database.html. Also, this database contains metadata of every abnormality the incorporation of the breast imaging reports and facts of system (BI-RADS) lexicon. Anomaly severity can be bifurcated into benign and malignant[9].
- INbreast INbreasthas a complete of a hundred and fifteen cases (410 pix) from which 90 instances are from ladies with each breasts affected (4 images in line with case) and 25 instances are from mastectomy patients (images in keeping with case).

V. DATA PRE-PROCESSING

In this stage initially, the part of the images that are unimportant or don't contain any region of interest are trimmed. Seeing that mammograms are taken under distinct conditions, they may be laid low with noise and a few artifacts. moreover, they generally do no longer have the favoured evaluation to carry out accurate analyses of the two proposed techniques. As such, the neighborhood vicinity histogram equalization is used and then the median filtering is carried out to lessen noise. In the histogram equalization degree, the magnitude of image pixels are expanded in order to extend the evaluation. Median filtering can be defined as a nonlinear operation that is frequently utilized in photograph processing to lessen salt and pepper and speckle noise[9].

VI. SEGMENTATION

The segmentation system trims off the regions of interest from the initial tissue in mammograms. The fundamental techniques in segmentation are: (i) region-based methods (which include vicinity developing, split/merge the usage of quad-tree decomposition) wherein similarities are detected, and (ii) boundary-based totally methods (including thresholding, gradient part detection) wherein discontinuities are detected and linked to shape area boundaries[9]. The segmentation of nontrivial pix is one of the maximum hard

obligations in photo processing. one of the strategies for segmenting tumors are 1)place developing method 2)mobile Neural network.
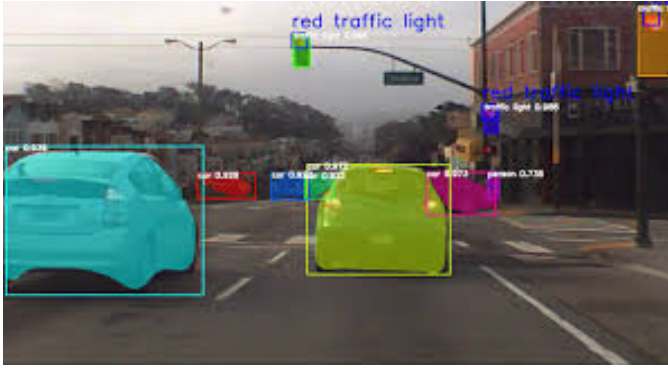

Fig. 2.   Segmentation used in object detection [11]

VII. INCEPTION-RESNET-V2

Inception-ResNet-v2 is a convolutional neural community (CNN) that achieves a brand new country of the artwork in terms of accuracy on the ILSVRC photograph type benchmark. Inception-ResNet-v2 is a version of the earlier Inception V3 version which borrows a few ideas from Microsoft's ResNet papers[12]. Very deep convolutional networks have been principal to the largest advances in picture reputation performance in recent years. One instance is the Inception structure that has been shown to reap excellent overall performance at incredibly low computational value. The creation of residual connections along with a greater conventional architecture has yielded ultra-modern performance inside the ILSVRC challenge; its overall performance became just like the contemporary era Inception-v3 community[18].

Model	Architecture	Checkpoint	Top-1 Accuracy	Top-5 Accuracy
Inception-ResNet-v2	Code	inception_resnet_v2_2016_08_30.tar.gz	80.4	95.3
Inception V3	Code	inception_v3_2016_08_28.tar.gz	78.0	93.9
ResNet 152	Code	resnet_v1_152_2016_08_28.tar.gz	76.8	93.2
ResNet V2 200	Code	TBA	79.9*	95.2*

(*): Results quoted in ResNet paper.

Fig. 3.   (*): Results quoted in ResNet paper.[12]

This increases the question of whether there are any advantage in combining the Inception structure with residual connections. there is clean empirical evidence that schooling with residual connections hurries up the training of Inception networks substantially. there is also a few proof of residual Inception networks outperforming similarly highly-priced Inception networks with out residual connections by means of a thin margin[18]. Residual connections permit shortcuts inside

the version and feature allowed researchers to efficiently teach even deeper neural networks, which have result in even higher performance. This has additionally enabled sizeable simplification of the Inception blocks[12].
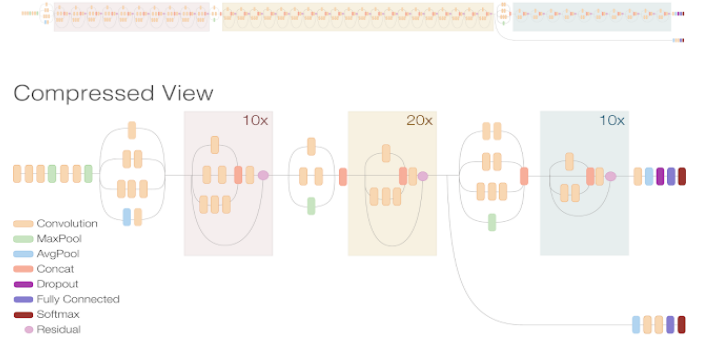

Fig. 4.   Schematic diagram of Inception-ResNet-v2 [12]

VIII. FEATURE EXTRACTION

In comparison to abnormal regions in mammogram reports, normal or benign regions seem to show lesser depth. Shape features have an effect on harmful and harmless differentiation energy; because abnormalities/tumors belonging to the identical tissue are of similar form. As harmful tumors frequently have erratic texture as compared to harmless tumors, textural capabilities are extracted from grey-stage co-occurrence matrix (GLCM) that have the order of statistical records of neighboring pixels of a photo. Amongst extracted features, Zernike moments are proper descriptors for item form.For extracting properties,any specific marginal data is no longer required. Even if the items are not segmented very well, they are able to reap top outcomes. Zernike moments map an photograph to a fixed of Zernike complex polynomials. When you consider that Zernike polynomials are orthogonal to every different, they gift picture functions without overlapping and further facts. The procedure of calculating Zernike moments associated with an photo are explained as follows:

- Calculate radius polynomials.
- Calculate Zernike basic functions.
- Map image matrix on Zernike basic functions to obtain Zernike Moments[9].

IX. FEATURE SELECTION

Characteristic selection is carried out for you to select appropriate capabilities from extracted functions. It improves the prediction preciseness and decreases the computational cost . Function choice is a seek trouble in a huge scope of panacea (unique combos of capabilities). The decision of the abilities, of the genetic algorithms are used with one of a kind kinds of chromosomes advent and health capabilities. Within the first shape, every chromosome is a binary string wherein each gene shows the presence or absence of every feature, as zero and 1. Within the second structure, every chromosome has twenty genes and each gene is assigned

with values accounting between one to fifty-one so as to selecting a single out of the myriad of 51 functions. The comprehensive steps for the genetic algorithm are given as:

- Obtain the embryonic population of the chromosomes.
- Reiterate.
- Calculate fitness function corresponding to each element in population (individual).
- Select pairs of the best ranking chromosomes as parents.
- Solicit the operator for cross over.
- Solicit the operator for mutation.
- Until the culminating condition.
- Stop[9].

X. CLASSIFICATION OF ABNORMALITY

Classification of abnormalities is based on a number of parameters. Some of which include mass density, architecture of distortion, calcification etc.
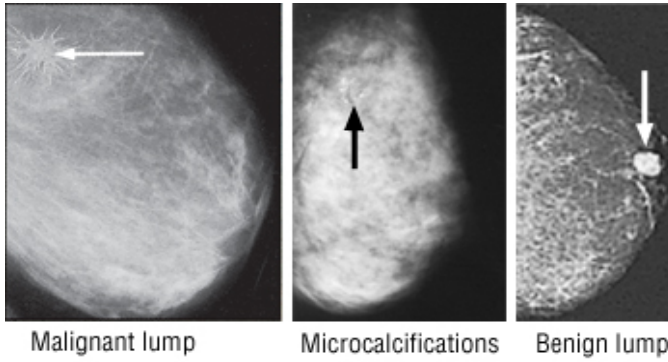

Fig. 5. Different types of lumps : Malignant and Benign along with Microcalcification [15]

A. BREAST IMAGINIG REPORTING AND DATA SYSTEM CATEGORIES

Category	Description
0	Incomplete; need additional imaging evaluation
1	Negative
2	Benign finding
3	Probably benign finding
4	Suspicious abnormality
5	Highly suggestive of malignancy
6	Known malignancy

BI-RADS ASSESSMENT CATEGORIES [16]

XI. PERFORMANCE DIFFERENTIATION

Error calculation of the used Multilayer Perceptron (MLP) neural networks is done by obtaining the mean squared error as follows:-
where O and F are the target and output matrices, respectively. other related metrics are also calculated as:
TP: actual superb, the category result is nice in presence of malignancy. TN: proper terrible, the class end result is

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(O_{(i)} - F_{(i)}\right)^2$$

Fig. 6. Errors of the used MLP neural networks are calculated by mean squared error (MSE) according to where O and F are the target and output matrices, respectively. [9]

terrible in being benign. FP: false positive, the category end result is superb in being benign. FN: fake negative, the class end result is negative in presence of malignancy. consistent with above definitions the equations associated with specificity (accuracy of poor elegance), sensitivity (accuracy of high quality class) and accuracy of understand both bad and high quality training are defined as[9]
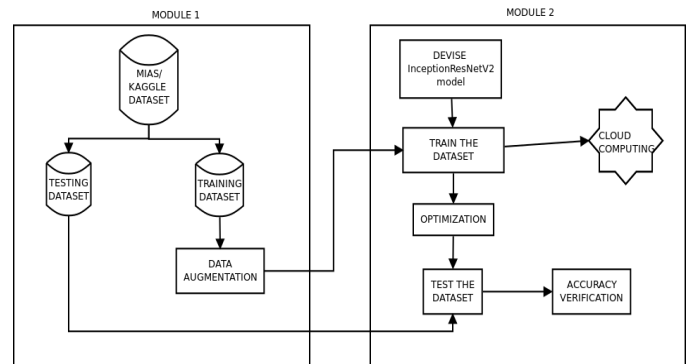
$$Specificity = \frac{TN}{TN + FP}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Fig. 7. Specificity, Sensitivity and Accuracy cab be evaluated by: [9]
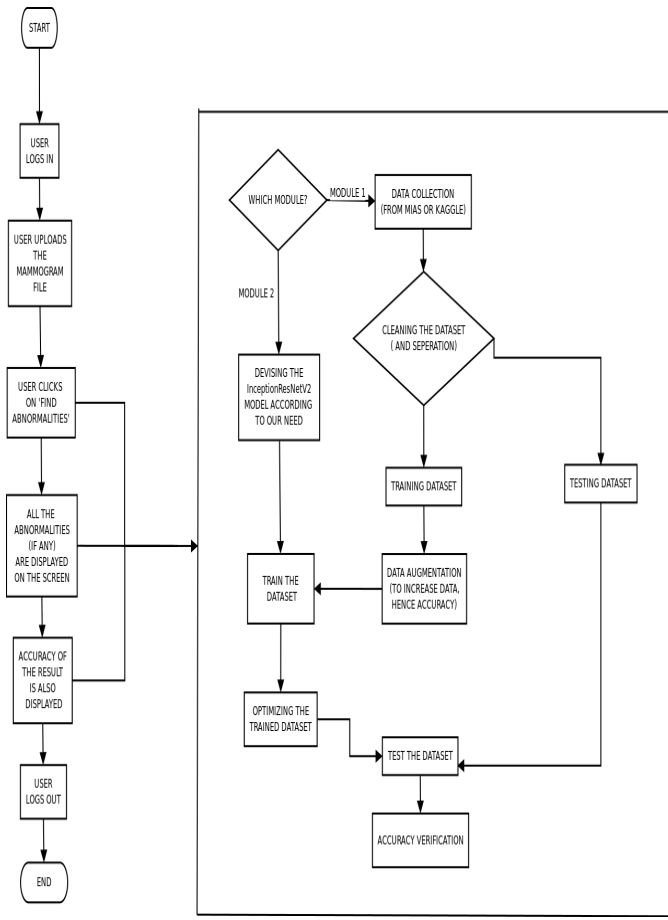
PROPOSED SYSTEM ARCHITECTURE

START → USER LOGS IN → USER UPLOADS THE MAMMOGRAM FILE → USER CLICKS ON 'FIND ABNORMALITIES' → ALL THE ABNORMALITIES (IF ANY) ARE DISPLAYED ON THE SCREEN → ACCURACY OF THE RESULT IS ALSO DISPLAYED → USER LOGS OUT → END

WHICH MODULE?

MODULE 1 → DATA COLLECTION (FROM MIAS OR KAGGLE) → CLEANING THE DATASET (AND SEPERATION) → TRAINING DATASET / TESTING DATASET

MODULE 2 → DEVISING THE InceptionResNetV2 MODEL ACCORDING TO OUR NEED

TRAIN THE DATASET ← DATA AUGMENTATION (TO INCREASE DATA, HENCE ACCURACY)

OPTIMIZING THE TRAINED DATASET → TEST THE DATASET → ACCURACY VERIFICATION

Fig. 8. Flowchart of the proposed system

XII. LITERATURE SURVEY

Sr.no.	Technology Used	Gap identified
1	Approximately 70 percent of the data set was trained on SGDM. Softmax layer is used in the fully connected layer for binary classification.	Mammogram patches were used to present augmented data set on which contrast enhancement was applied. Images were resized from 1024 pixels to 224 pixels. Data was trained randomly to yield better results.
2	Project was implemented on Tensorflow. Hidden layer in CNN consisted of convolutional layer, ReLU and fully connected dense layer.	Input images were cropped to 48 pixels. To increase accuracy images were transformed. Abnormality tissues that are too close are removed.
3	Classification based on Softmax regression and SVM to generate probability distribution function. Activation functions include Stigmoid, tanh and ReLU	Research presents a simple network structure for image classification with small memory and good recognition effect.
4	Back propogation	The main aim is procuring the unparalleled activation function that reduces the classification error. The research presents logsigmoid function and hyperbole tangent without biases as a low complexity architecture for Back Propagation.
5	All experiments conducted during the research were implemented with Caffe.Batch normalization was used on input images consisting of 3 convolution layers. Activation functions used are ReLU, Xavier weight initialization and Adam.	Three different architectures were used: Shallow CNN, Alexnet and GoogleNet. Research implies that area surrounding the mass provides useful context for diagnosis, where proportionally large padding contains greater signal for classification. From the above models, GoogleNet is least prone to overfitting.
6	Study detects cancer tumour using K nearest neighbour algorithm. Matlab is used for implementation along with Weiner filter.	Image processing techniques along with machine learning algorithms were used to transform the images from time domain to frequency domain using Discrete Wavelet Transformation.
7	Study makes use of Gabor filter, SVM and MLP classifier. Input patches were normalized using Gaussian pyramid processing. Model is trained on Stochastic Gradient Descent.	Geometric transformation was applied to obtain a large data set of mammogram images. Input was normalized and scaled using Gaussian pyramid. Model learns in a hierarchical manner.

8	Research is implemented using conventional, region based and feature based techniques.Two types of segmentation used are single view mass detection and multiple view mass detection.	Research focuses a lot on the medical aspect of the project.
9	This paper uses Cellular Neural Network, Region Growing segmentation method, Genetic Algorithm, Artificial Neural Network,SVM.	In this work, 2 automated methods were presented based on the improvement of region growing and CNN segmentation to obtain an self changing threshold and acceptable templates, respectively, in order to preserve tumor boundary information to diagnose benign and malignancy in mammograms.
10	This paper uses Particle Swarm Optimized Wavelet Neural Network (PSOWNN), Wavelet neural network (WNN), Receiver Operating Characteristic (ROC) curve, Particle swarm optimization (PSO).	The model approach demonstrated in this paper stated that the PSOWNN classifier produces an improvement in differentiation proficiency to the issue of computer-aided analysis of digital mammograms for abnormality detection in the mammogram report .Good differentiation proficiency in WNN based classifier is achieved by minimizing the amount of false positives and negatives.The WNN classifier that are under consideration use the properties of both wavelet and neural network.

A. ADVANTAGES

- The proposed system is the first of it's kind to incorporate InceptionResnetV2 as the base model to diagnose mammograms.
- The proposed system will a public and free website to diagnose the parameters leading to breast cancer.
- Due to good accuracy of computer aided detectors (CAD) the patients have a good clarity of their medical report.

B. DISADVANTAGES

- A distorted image can yield false results.

XIII. CONCLUSION

The study observed and analyzed the different abnormalities in a mammogram and the implementation of various strategies to determine the abnormalities successfully in a mammogram. Analyzing various models, we have come to a conclusion that the InceptionResnetV2 model is one of the best models for processing images (mammograms) which would greatly assist in the detection of abnormalities. This system being free and open to the public, will greatly help the patients in having a quicker, cheaper and accurate diagnosis of mammograms.

FUTURE SCOPE

This system focuses on the early detection of breast cancer with the help of CNN. Earlier the detection, higher is the safety. Therefore, the future scope of this project would be to incorporate better image processing techniques and deeper neural networks along with a very large dataset of mammograms. This would help the system to detect tumour much earlier than other systems.

ACKNOWLEDGMENT

REFERENCES

[1] Charan, S., Khan, M. J., Khurshid, K. (2018, March). Breast cancer detection in mammograms using convolutional neural network. In Computing, Mathematics and Engineering Technologies (iCoMET), 2018 International Conference on (pp. 1-5). IEEE.

[2] Tan, Y. J., Sim, K. S., Ting, F. F. (2017, November). Breast cancer detection using convolutional neural networks for mammogram imaging system. In Robotics, Automation and Sciences (ICORAS), 2017 International Conference on (pp. 1-5). IEEE.

[3] Guo, T., Dong, J., Li, H., Gao, Y. (2017, March). Simple convolutional neural network on image classification. In Big Data Analysis (ICBDA), 2017 IEEE 2nd International Conference on (pp. 721-724). IEEE.

[4] Jouni, H., Issa, M., Harb, A., Jacquemod, G., Leduc, Y. (2016, November). Neural Network architecture for breast cancer detection and classification. In Multidisciplinary Conference on Engineering Technology (IMCET), IEEE International (pp. 37-41). IEEE.

[5] Lvy, D., Jain, A. (2016). Breast mass classification from mammograms using deep convolutional neural networks. arXiv preprint arXiv:1612.00542.

[6] Alarabeyyat, A., Alhanahnah, M. (2016, August). Breast Cancer Detection Using K-Nearest Neighbor Machine Learning Algorithm. In 2016 9th International Conference on Developments in eSystems Engineering (DeSE) (pp. 35-39). IEEE.

[7] Bakkouri, I., Afdel, K. (2017, May). Breast tumor classification based on deep convolutional neural networks. In Advanced Technologies for Signal and Image Processing (ATSIP), 2017 International Conference on (pp. 1-6). IEEE.

[8] Hela, B., Hela, M., Kamel, H., Sana, B., Najla, M. (2013, March). Breast cancer detection: A review on mammograms analysis techniques. In Systems, Signals Devices (SSD), 2013 10th International Multi-Conference on (pp. 1-6). IEEE.

[9] Rouhi, R., Jafari, M., Kasaei, S., Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Systems with Applications, 42(3), 990-1002.

[10] Dheeba, J., Singh, N. A., Selvi, S. T. (2014). Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. Journal of biomedical informatics, 49, 45-52.

[11] Google Search—https://www.google.com/imgres?imgurl=https

[12] Google AI Blog—Improving Inception and Image Classification in TensorFlow— August 24, 2018

[13] Journal of Surgery— http://www.avensonline.org/fulltextarticles/JSUR-2332-4139-S1-0001.html—August 25,2018

[14] Journal of Cancer Science Therapy—http://www.omicsonline.org/open-access/natural-history-of-ovarian-cancer-1948-5956.1000278.php?aid=27726— August 25,2018

[15] Journal Title Journal of Prosthodontics —Biological Aspects as a Rule for Single Implant Placement. The 3A-2B Rule: A Clinical Report— August 27, 2018

[16] Clinical Breast Tomosynthesis— 2 BI-RADS nomenclatureBI-RADS Nomenclature for Mammography and Ultrasound— August 26, 2018

[17] Cancer Nursing— Age Differences in Post???Breast Cancer Lym-phedema Signs and Symptoms— September 8, 2018

[18] arXiv:1602.07261 [cs.CV]